Amendments to the Claims:

This listing of claims will replace all prior versions, and listings of claims in the application. Applicant has submitted a new complete claim set showing marked up claims with insertions indicated by underlining and deletions indicated by strikeouts and/or double bracketing.

Listing of Claims:

1. (Currently amended) A method comprising:

selecting a modeling parameter from a plurality of modeling parameters characterizing a mixture of Student distribution components;

computing an approximation of a posterior distribution for the selected modeling parameter based on an input set of data<u>, the input set of data having been obtained from at least one microphone,</u> and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters, computing the approximation being performed by a processor calculating $q(\mathbf{s})=\prod_{n,m}^{N,M} p_{nm}^{s_{nm}}$, $q(\pi)=D\left(\pi|\alpha\right)$, $q(\boldsymbol{\mu}_m)=N\left(\boldsymbol{\mu}_m|m_m,\mathbf{R}_m\right)$, $q(\boldsymbol{\Lambda}_m)=W\left(\boldsymbol{\Lambda}_m|_m\mathbf{W}_m,\eta_m\right)$, or $q(u_{nm})=G\left(u_{nm}|a_{nm},b_{nm}\right)$;

computing a lower bound of a log marginal likelihood as a function of current estimates of the posterior distributions of the modeling parameters, the current estimates of the posterior distributions of the modeling parameters including the computed approximation of the posterior distribution of the selected modeling parameter;

determining if the lower bound has been satisfactorily optimized, wherein the lower bound is satisfactorily optimized when the computed lower bound has changed less than a threshold amount from a previously computed lower bound;

generating a probability density modeling the input set of data, the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters, when the lower bound is satisfactorily optimized;

outputting the probability density, and

outputting a number of speakers from the probability density.

2. (Previously presented) The method of claim 1 wherein the computing operations comprise a first iteration and further comprising:

selecting a different modeling parameter from the plurality of modeling parameters and repeating in a subsequent iteration the operations of computing an approximation and computing a lower bound using the newly selected modeling parameter, when the lower bound is not satisfactorily optimized in the first iteration.

3. (Original) The method of claim 1 wherein computing a lower bound comprises:

computing the lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters.

4. (Previously presented) The method of claim 1 wherein computing an approximation of a posterior distribution comprises:

computing a variational approximation of the posterior distribution of the selected modeling parameter.

5. (Original) The method of claim 1 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

6. (Original) The method of claim 1 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

7. (Cancelled).

8. (Original) The method of claim 1 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

9. (Original) The method of claim 1 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

10. (Original) The method of claim 1 wherein generating a probability density comprises:

generating the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the

current estimates of the posterior distributions of the modeling parameters and an estimate of the number of degrees of freedom of each Student distribution component.

11. (Original) The method of claim 1 further comprising:

storing the current estimates of the posterior distributions of the modeling parameters in a storage location.

12. (Previously presented) The method of claim 1 wherein the input set of data represents auditory speech data from an unknown number of speakers.

13. (Canceled).

14. (Currently amended) A computer program product encoding a computer program for executing on a computer system a computer process for minimizing effects of outlier data on data modeling, the computer process comprising:

selecting a modeling parameter from a plurality of modeling parameters characterizing a mixture of Student distribution components;

computing an approximation of a posterior distribution for the selected modeling parameter based on an input set of data<u>, the input set of data having been obtained from at least one microphone,</u> and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters, the current estimate being computed using $p(\mathbf{s}|\pi)=\prod_{n,m}^{N,M}\pi_m^{s_{nm}}$,

$p(\boldsymbol{\mu}_m)=N\ (\boldsymbol{\mu}_m|\mathbf{m},\rho\mathbf{I})$, $p(\boldsymbol{\Lambda}_m)=W\ (\boldsymbol{\Lambda}_m|\mathbf{W}_0,\eta_0)$, or $p(\pi)=D\ (\pi|\boldsymbol{\alpha})$;

computing a lower bound of a log marginal likelihood as a function of current estimates of the posterior distributions of the modeling parameters, the current

estimates of the posterior distributions of the modeling parameters including the computed approximation of the posterior distribution of the selected modeling parameter;

determining if the lower bound has been satisfactorily optimized, wherein the lower bound is satisfactorily optimized when the computed lower bound has changed less than a threshold amount from a previously computed lower bound;

generating a probability density modeling the input set of data, the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters, when the lower bound is satisfactorily optimized;

outputting the probability density; and

outputting a number of ~~clusters~~ speakers from the probability density.

15. (Previously presented) The computer program product of claim 14 wherein the computing operations comprise a first iteration and further comprising:

selecting a different modeling parameter from the plurality of modeling parameters and repeating in a subsequent iteration the operations of computing an approximation and computing a lower bound using the newly selected modeling parameter, when the lower bound is not satisfactorily optimized in the first iteration.

16. (Original) The computer program product of claim 14 wherein computing a lower bound comprises:

computing the lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters.

17. (Previously presented) The computer program product of claim 14 wherein computing a approximation of a posterior distribution comprises:

computing a variational approximation of the posterior distribution of the selected modeling parameter.

18. (Original) The computer program product of claim 14 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

19. (Original) The computer program product of claim 14 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

20. (Cancelled).

21. (Original) The computer program product of claim 14 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

22. (Original) The computer program product of claim 14 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

23. (Original) The computer program product of claim 14 wherein generating a probability density comprises:

generating the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters and an estimate of the degrees of freedom of each Student distribution component.

24. (Original) The computer program product of claim 14 wherein the computer process further comprises:

storing the current estimates of the posterior distributions of the modeling parameters in a storage location.

25. (Currently Amended) The computer program product of claim 14 wherein the input set of data represents auditory speech data from an unknown number of speakers~~, and wherein the number of clusters corresponds to the unknown number of speakers~~.

26. (Previously presented) The computer program product of claim 14 wherein the input set of data represents image segmentation data from images.

27. (Currently amended) A system comprising:

a processor;

a memory;

<u>at least one microphone;</u>

a modeling parameter selector operable with the processor and memory to select a modeling parameter from a plurality of modeling parameters characterizing a mixture of Student distribution components;

an approximation module computing an approximation of a posterior distribution for the selected modeling parameter based on an input set of data, the input set of data having been obtained from the at least one microphone, and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters;

a lower bound optimizer module computing a lower bound of a log marginal likelihood as a function of current estimates of the posterior distributions of the modeling parameters using $L\ (\boldsymbol{q}) \equiv \int q(\theta) \ln \left\{ \frac{p(\mathbf{X},\theta)}{q(\theta)} \right\} d\theta \leq \ln p(\mathbf{X})$, the current estimates of the posterior distributions of the modeling parameters including the computed approximation of the posterior distribution of the selected modeling parameter, and determining if the lower bound has been satisfactorily optimized, wherein the lower bound is satisfactorily optimized when the computed lower bound has changed less than a threshold amount from a previously computed lower bound;

a data model generator generating a probability density modeling the input set of data, the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters, when the lower bound is satisfactorily optimized;

an output device outputting the probability density and outputting a number of ~~clusters~~ speakers from the probability density.

28. (Original) The system of claim 27 wherein the lower bound optimizer computes the lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters.

29. (Previously presented) The system of claim 27 wherein the approximation module computes a variational approximation of the posterior distribution of the selected modeling parameter.

30. (Original) The system of claim 27 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

31. (Original) The system of claim 27 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

32. (Cancelled).

33. (Original) The system of claim 27 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

34. (Original) The system of claim 27 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

35. (Original) The system of claim 27 wherein the data model generator generates the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters and an estimate of the degrees of freedom of each Student distribution component.

36. (Original) The system of claim 27 further comprising:

a memory storing the current estimates of the posterior distributions of the modeling parameters.

37. (Currently Amended) The system of claim 27 wherein the input set of data represents auditory speech data from an unknown number of speakers~~, and wherein the number of clusters corresponds to the unknown number of speakers~~.

38. (Previously presented) The system of claim 27 wherein the input set of data represents image segmentation data from images.

39. (Currently amended) A method comprising:

computing an approximation of a posterior distribution for a selected modeling parameter of a plurality of modeling parameters characterizing a mixture of Student distribution components based on an input set of data<u>, the input set of data having been obtained from a microphone,</u> and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters, wherein computing the approximation is performed by a processor calculating

$$q(\mathbf{s}) = \prod_{n,m}^{N,M} p_{nm}^{s_{nm}}, \; q(\pi) = D\left(\pi \middle| \alpha\right), \; q(\boldsymbol{\mu}_m) = N\left(\boldsymbol{\mu}_m \middle| m_m, \mathbf{R}_m\right), \; q(\boldsymbol{\Lambda}_m) = W\left(\boldsymbol{\Lambda}_m \middle|_m \mathbf{W}_m, \eta_m\right),$$

or $q(u_{nm}) = G\left(u_{nm} \middle| a_{nm}, b_{nm}\right)$;

determining whether current estimates of the posterior distributions of the modeling parameters are satisfactorily optimized in relation to a predetermined criterion, the current estimates of the posterior distributions of the modeling parameters including the computed approximation of the posterior distribution of the selected modeling parameter;

modeling the input set of data by the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters;

outputting the modeling of the input set of data; and

outputting a number of ~~clusters~~ speakers from the probability density.

40. (Previously presented) The method of claim 39 wherein the computing operation and determining operation comprise a first iteration and further comprising:

selecting a different modeling parameter from the plurality of modeling parameters and repeating in a subsequent iteration the operations of computing a approximation and computing a lower bound using the newly selected modeling parameter, when the lower bound is not satisfactorily optimized in the first iteration.

41. (Previously presented) The method of claim 39 wherein the operation of determining whether current estimates of the posterior distributions of the modeling parameters are satisfactorily optimized comprises:

computing a lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters and a variational posterior distribution; and

determining whether the lower bound satisfies the predetermined criterion of the selected modeling parameter.

42. (Previously presented) The method of claim 39 wherein computing a approximation of a posterior distribution comprises:

computing a variational approximation of the posterior distribution.

43. (Original) The method of claim 39 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

44. (Original) The method of claim 39 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

45. (Cancelled).

46. (Original) The method of claim 39 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

47. (Original) The method of claim 39 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

48. (Original) The method of claim 39 wherein modeling the input data comprises:

generating the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters and an estimate of the degrees of freedom of each Student distribution component.

49. (Original) The method of claim 39 further comprising:

storing the current estimates of the posterior distributions of the modeling parameters in a storage location.

50. (Currently amended) A computer program product encoding a computer program for executing on a computer system a computer process for minimizing effects of outlier data on data modeling, the computer process comprising:

computing an approximation of a posterior distribution for a selected modeling parameter of a plurality of modeling parameters characterizing a mixture of Student distribution components based on an input set of data, the input set of data having been obtained from a microphone, and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters, computing the approximation being performed by a processor calculating

$$q(\mathbf{s}) = \prod_{n,m}^{N,M} p_{nm}^{s_{nm}},\ q(\pi) = D\left(\pi \middle| \alpha\right),\ q(\boldsymbol{\mu}_m) = N\ \left(\boldsymbol{\mu}_m \middle| m_m, \mathbf{R}_m\right),\ q(\boldsymbol{\Lambda}_m) = W\ \left(\boldsymbol{\Lambda}_m \middle|_m \mathbf{W}_m, \eta_m\right),$$

or $q(u_{nm}) = G\left(u_{nm} \middle| a_{nm}, b_{nm}\right)$;

determining whether current estimates of the posterior distributions of the modeling parameters are satisfactorily optimized in relation to a predetermined criterion, the current estimates of the posterior distributions of the modeling parameters including the computed approximation of the posterior distribution of the selected modeling parameter;

modeling the input set of data by the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters;

outputting the modeling of the input set of data; and

outputting a number of ~~clusters~~ speakers from the probability density.

51. (Previously presented) The computer program product of claim 50 wherein the computing operation and determining operation comprise a first iteration and further comprising:

selecting a different modeling parameter from the plurality of modeling parameters and repeating in a subsequent iteration the operations of computing a approximation and computing a lower bound using the newly selected modeling parameter, when the lower bound is not satisfactorily optimized in the first iteration.

52. (Previously presented) The computer program product of claim 50 wherein the operation of determining whether current estimates of the posterior distributions of the modeling parameters are satisfactorily optimized comprises:

computing a lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters and a variational posterior distribution; and

determining whether the lower bound satisfies the predetermined criterion.

53. (Previously presented) The computer program product of claim 50 wherein computing an approximation of a posterior distribution comprises:

computing a variational approximation of the posterior distribution of the selected modeling parameter.

54. (Original) The computer program product of claim 50 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

55. (Original) The computer program product of claim 50 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

56. (Cancelled).

57. (Original) The computer program product of claim 50 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

58. (Original) The computer program product of claim 50 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

59. (Original) The computer program product of claim 50 wherein modeling the input data comprises:

generating the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters and an estimate of the degrees of freedom of each Student distribution component.

60. (Original) The computer program product of claim 50 wherein the computer process further comprises:

storing the current estimates of the posterior distributions of the modeling parameters in a storage location.

61. (Currently amended) A system comprising:

a processor;

a memory;

at least one microphone;

a approximation module operable with the processor and memory to compute an approximation of a posterior distribution for a selected modeling parameter of a plurality of modeling parameters characterizing a mixture of Student distribution components based on an input set of data, the input set of data having been obtained form the at least one microphone, and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters by calculating $q(\mathbf{s}) = \prod_{n,m}^{N,M} p_{nm}^{s_{nm}}$, $q(\pi) = D\left(\pi|\alpha\right)$, $q(\boldsymbol{\mu}_m) = N\left(\boldsymbol{\mu}_m|m_m, \mathbf{R}_m\right)$, $q(\boldsymbol{\Lambda}_m) = W\left(\boldsymbol{\Lambda}_m|_m \mathbf{W}_m, \eta_m\right)$, or $q(u_{nm}) = G\left(u_{nm}|a_{nm}, b_{nm}\right)$;

an optimizer module determining whether current estimates of the posterior distributions of the modeling parameters are satisfactorily optimized in relation to a predetermined criterion, the current estimates of the posterior distributions of the modeling parameters including the computed approximation of the posterior distribution of the selected modeling parameter;

a data model generator modeling the input set of data by the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters;

an output device outputting the modeling of the input set of data and outputting a number of ~~clusters~~ speakers from the probability density.

62. (Previously presented) The system of claim 61 wherein optimizer module computes a lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters and a variational posterior distribution, and determines whether the lower bound satisfies the predetermined criterion.

63. (Previously presented) The system of claim 61 wherein the approximation modules computes a variational approximation of the posterior distribution of the selected modeling parameter.

64. (Original) The system of claim 61 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

65. (Original) The system of claim 61 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

66. (Cancelled).

67. (Original) The system of claim 61 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

68. (Original) The system of claim 61 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

69. (Original) The system of claim 61 wherein modeling the input data comprises:

generating the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters and an estimate of the degrees of freedom of each Student distribution component.

70. (Original) The system of claim 61 further comprising:

a memory storing the current estimates of the posterior distributions of the modeling parameters.

71. (Previously presented) The method of claim 1 further comprising populating the input set of data with only observed data.